Demonstration of a 3D Vision Algorithm for Space Applications

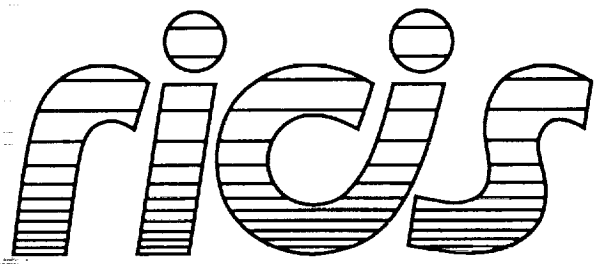
Rui J. P. deFigueiredo, ed.

Rice University

1987

Cooperative Agreement NCC 9-16
Research Activity No. AI.05

NASA Johnson Space Center
Mission Planning and Analysis Division



Research Institute for Computing and Information Systems
University of Houston-Clear Lake

TECHNICAL REPORT

The RICIS Concept

The University of Houston-Clear Lake established the Research Institute for Computing and Information Systems (RICIS) in 1986 to encourage the NASA Johnson Space Center (JSC) and local industry to actively support research in the computing and information sciences. As part of this endeavor, UHCL proposed a partnership with JSC to jointly define and manage an integrated program of research in advanced data processing technology needed for JSC's main missions, including administrative, engineering and science responsibilities. JSC agreed and entered into a continuing cooperative agreement with UHCL beginning in May 1986, to jointly plan and execute such research through RICIS. Additionally, under Cooperative Agreement NCC 9-16, computing and educational facilities are shared by the two institutions to conduct the research.

The UHCL/RICIS mission is to conduct, coordinate, and disseminate research and professional level education in computing and information systems to serve the needs of the government, industry, community and academia. RICIS combines resources of UHCL and its gateway affiliates to research and develop materials, prototypes and publications on topics of mutual interest to its sponsors and researchers. Within UHCL, the mission is being implemented through interdisciplinary involvement of faculty and students from each of the four schools: Business and Public Administration, Education, Human Sciences and Humanities, and Natural and Applied Sciences. RICIS also collaborates with industry in a companion program. This program is focused on serving the research and advanced development needs of industry.

Moreover, UHCL established relationships with other universities and research organizations, having common research interests, to provide additional sources of expertise to conduct needed research. For example, UHCL has entered into a special partnership with Texas A&M University to help oversee RICIS research and education programs, while other research organizations are involved via the "gateway" concept.

A major role of RICIS then is to find the best match of sponsors, researchers and research objectives to advance knowledge in the computing and information sciences. RICIS, working jointly with its sponsors, advises on research needs, recommends principals for conducting the research, provides technical and administrative support to coordinate the research and integrates technical results into the goals of UHCL, NASA/JSC and industry.

Demonstration of a 3D Vision Algorithm for Space Applications

RICIS Preface

This research was conducted under auspices of the Research Institute for Computing and Information Systems by Dr. Rui J. P. deFigueiredo of Rice University. Dr. Terry Feagin served as RICIS research coordinator.

Funding was provided by the Mission Planning and Analysis Division, NASA/JSC through Cooperative Agreement NCC 9-16 between the NASA Johnson Space Center and the University of Houston-Clear Lake. The NASA research coordinator for this activity was Dr. Timothy F. Cleghorn, NASA/JSC.

Reprint from the Proceedings of the Space Telerobotics Workshop held at the Pasadena Center, Pasadena, CA, January 20-22, 1987

A Technique for 3D Robot Vision for Space Applications[1]

Vishal Markandey, Hemant Tagare, and Rui J.P. deFigueiredo

Dept of Electrical and Computer Engineering

Rice University, Houston, Texas 77251-1892

1) ABSTRACT:

This paper reports an extension of the MIAG algorithm for recognition and motion parameter determination of general 3D polyhedral objects based on model matching techniques and using *Moment Invariants* as features of object representation. Results of tests conducted on the algorithm under conditions simulating space conditions are presented.

2) INTRODUCTION:

Many different object recognition and attitude determination techniques have been proposed by researchers. The earliest ones used the approach of matching the observed image to a library of a fixed number of views of objects. The limitations of such an approach are glaringly apparent. Among the later techniques, Richard and Hemami [1] used Fourier descriptors and Dudani et al [2] used moment invariants. Watson and Shapiro [3] used a model matching technique to identify wireframe perspective views. Their method is iterative and requires use of a numerical optimization technique. Marr and Poggio [4] have implemented a stereo reconstruction algorithm which uses geometric constraints to recover surface shape. Similar range data techniques have been developed by other researchers also. A fundamental limitation of these techniques is the introduction of restrictive assumptions about the imaged scene in terms of generalized cones [5] or in terms of planar and quadric patches [6]. Horn [7] has worked on the extraction of shape from shading, using the reflectance map. This method uses the brightness gradient as the image feature used in recovery. It is applicable to smooth, uniform Lambertian surfaces. Stevens [8], Kender [9] and later Witkin [10] have tried to recover shape from texture. This technique and also the shape from contour (surface boundaries) technique presented by Barrow and Tenenbaum [11] rely on the assumption that the world of objects is regular. Such techniques are limited to smooth-textured surfaces. For some other contributions see Silcox [12].

Bamieh and deFigueiredo [12] have developed the Moment-Invariants / Attributed Graph (MIAG) Algorithm in which 2D moment invariants which are invariant under 3D motion, have been used for the recognition of 3D objects, using an attributed graph representation and based on the concept of model matching. This approach avoids restrictive geometric assumptions and so offers an advantage over most techniques discussed above. In its original form this algorithm was applicable for recognition of polyhedral objects but it could not be used for attitude determination if the polyhedron had symmetric faces for reasons discussed later in this paper. This limitation has been overcome now as discussed in this paper. As this technique uses moment invariants as features of representation and these can be computed only for planar faces, the technique is not directly applicable to the recognition and attitude determination of curved objects. However we can use other representational features such as the Gaussian and mean curvature and the attributed graph and model matching techniques can still be applied.

To implement this technique we need picture information in the form of wireframes. So the picture from the camera is digitized and converted to wireframe form before applying the MIAG algorithm to it. In the work currently in progress, the overall process is divided into three parts:

[1]Supported by the NASA contract NAS 9-17145, the NASA/JSC grant NCC 9-16, and the NSF grant DCR-8318514.

1) Data acquisition and digitization.

2) Wireframe extraction.

3) Recognition and motion parameter determination.

The first two parts above constitute the image processing/feature extraction stage. The work in this stage is briefly outlined below.

3) IMAGE PROCESSING / FEATURE EXTRACTION:

3.1) Data acquisition and digitization: Models of various space objects, such as mockups of satellites, the space shuttle and parts of the space station are being used to test the performance of the algorithm. These models are grabbed by cameras under illumination conditions simulating those prevelent in space orbit. The pictures are digitized to obtain 2D arrays of brightness values. This is the initial level of representation in the system.

3.2) Wireframe extraction: Wireframe extraction consists of removing noise from the picture and subjecting it to edge detection and reconstruction. The input image is lowpass filtered to remove the high frequency noise. A 7x7 Gaussian filter is used for this. The Sobel gradient operator is then applied to the output of the lowpass filter to obtain an edge detected version of the input image. This image is a grey level image. It is converted to binary form by thresholding, which also removes some of the noise and thins down the edges. The remaining noise is removed by median filtering. A length 5 filter was employed for this. The output so obtained is a noise free, binary edge image. But the edges are thick smears instead of the fine lines required in a wireframe. A thinning algorithm [13] is applied to this to reduce the edges to unit pixel thickness, thus obtaining the required wireframe.

4) RECOGNITION AND MOTION PARAMETER DETERMINATION:

The MIAG algorithm [12] is an algorithm of recognition and attitude determination of 3D objects. We discuss this algorithm in two steps: First, object recognition and second, attitude determination.

4.1) Object recognition: The MIAG technique recognizes a 3D object from its projection on an imaging plane. The algorithm works for the identification of polyhedral objects. Each face of a polyhedron can be considered as a rigid planar patch (RPP). Motion of the object can then be considered as motion of its constituent RPP's. If it is assumed that the image is formed by parallel projection then if an RPP undergoes rigid body motion in 3D its image undergoes affine transformations. So the method which tries to identify an object in 3D motion should use features of images which remain invariant under affine transformations. General moment invariants are such features. They remain invariant under translation, rotation and scale changing. Moments are coefficients in a series expansion of the image function, similar to those in a Fourier series expansion. But unlike in Fourier series where sine and cosine functions are the basis functions, here the basis functions are polynomials in the image function variables. Thus if the picture function is f(x,y) its moment is:

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy$$

for p,q=0,1,2...

The value of (p+q) is known as the order of the moment. Theoretically, for a perfect description of the picture in terms of moments, p and q should go to ∞. But in the present algorithm moments upto only order four have been used. This is because the computation of higher order moments is increasingly difficult and it was found that picture representation in terms of four moments gives good results in the test cases.

These moments have to be computed for each face of the picture wireframe. The picture intensity is taken to be 1 inside a polygon and 0 outside it. Thus all the picture information is contained in its boundaries. Using this fact, the above surface integral can be changed to a line integral by Green's theorem. For a digital picture the integral reduces to a sum. See [12] for details.

Moment invariants of all faces of certain standard objects are stored in the system library. Given a wireframe which needs to be identified, the moment invariants of each of its faces are computed. These are then matched to the stored values of the moment invariants of the library objects. If all the moments corresponding to a face of the RPP match all the moments of a face of a stored object, we can say that the two faces are similar. For the objects to be the same, not only should the faces be similar but the adjacency

conditions of the faces and the angles between the faces should be similar. To carry out this matching, the wireframe is first converted to an attributed graph. Each node of the graph represents a face of the wireframe. If two nodes are connected by a line (edge) it means that the faces corresponding to these nodes are adjacent. With each node is associated a feature vector consisting of a set of moment invariants of the face that it represents and with each edge is associated a scalar which gives the angle between normals to the two faces it connects. As an example, the attributed graph representation of a cube is shown in Fig 1.

The algorithm works as follows: Suppose we hypothesize that node W_j in the wireframe corresponds to node O_j in the model graph. If W_j has k nodes adjacent to it, $W_{m1},...,W_{mk}$ then O_j should also have k nodes adjacent to it, $O_{n1},...,O_{nk}$. The following constraints should be satisfied:

1) W_{ms} must have the same feature vectors as O_{ns}, s=1,...,k.

2) Angle between W_{ms} and W_j should equal angle between O_{ns} and O_j for all s=1,...,k.

3) If any two W_{ms} 's are connected then the angle between them should equal that between their matching nodes.

If all these conditions are satisfied, then an admissible matching configuration is said to have been obtained at nodes W_j and O_j. If matching configurations are obtained between all nodes of the given wireframe and one of the stored models, then we can say that the wireframe matches the model. In most cases, only a small part of the given model needs to be matched to discriminate it against the other models in the library.

It may be noted that because of numerical truncations and rounding during calculations there may not be a perfect match between the moment invariants computed for the wireframe and those stored for the model. So we define a measure of error between the two sets of moment invariants. The moment invariants can be taken as coordinates of a point in four dimensioinal vector space and the distance between the two points is taken as a measure of error. If I_1, I_2, I_3 and I_4 are the moment invariants of a wireframe's face and I'_1, I'_2, I'_3 and I_4 those of a model's face then the distance is:

$$d = \sqrt{(I_1 - I'_1)^2\rho_1^2 + (I_2 - I'_2)^2\rho_2^2 + (I_3 - I'_3)^2\rho_3^2 + (I_4 - I'_4)^2\rho_4^2}$$

where the ρ's are weighting factors. These are needed to equalize the contribution of all four moment invariants in the error measure because some of the moment invariants may have values of the order of 10^{-3} and others may be of the order of 10^{-7}. If the value 'd' is less than a certain threshold (taken 0.01 here), the two sets of moment invariants are taken to be equivalent.

The driver algorithm arbitrarily picks a node W_j in the wireframe, then it looks for a node O_j in the model with the same feature vector. If matched, these nodes are marked as a pair. An adjacent image face is chosen and the adjacent object faces are scanned to see if one of them matches it. As each adjacent pair is found it is checked for consistent adjacency and equality of angles between faces. If everything matches satisfactorily a succesful match is declared.

4.2) Attitude determination: The identity of the object having been so determined, one has to estimate the attitude and location of the recognized object relative to a library standard.

Let (X,Y,Z) be the original coordinates of a point on the body and (X',Y',Z') be its coordinates after motion. Then,

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T$$

where

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_8 \end{bmatrix}$$

is the rotation matrix and

$$T = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix}$$

is the translation matrix.

Let the corresponding points on the image be (x,y) and (x',y'). Then,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r_1 x + r_2 y + r_3 Z + \Delta x \\ r_4 x + r_5 y + r_6 Z + \Delta y \end{bmatrix}$$

For simplicity let Z=0 i.e. the RPP in library lies in the x-y plane of the object space. Then the above matrix can be repressented as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} Q_1^1 & Q_2^1 \\ Q_1^2 & Q_2^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \end{bmatrix}$$

where

$Q_1^1 = r_1$, $Q_2^1 = r_2$, $Q_1^2 = r_4$, $Q_2^2 = r_5$, $\Delta x = \Delta X$, $\Delta y = \Delta Y$.

All Q's and Δ's can be determined from the moments of the images. To find the rest of the r's we use the fact that the sum of squares of any row or column in the r matrix is 1.

From the r's we can find directional cosines of the rotation axis and the rotation angle as:

$$\sin\theta = \frac{d}{2} \text{ and } \cos\theta = \frac{d^2 r_1 - (r_8 - r_{6)}^2}{d^2 - (r_8 - r_{6)}^2}$$

where

$$d = \sqrt{(r_8 - r_6)^2 + (r_3 - r_7)^2 + (r_4 - r_2)^2}$$

(both sin θ and cos θ are needed to determine θ uniquely). The direction cosines are:

n1 = (r_8 - r_6) / d

n2 = (r_3 - r_7) / d

n3 = (r_4 - r_2) / d

where $d^2 = (r8-r6)^2 + (r3-r7)^2 + (r4-r2)^2$

We can also find the translation in x and y directions as:

$\Delta x = m_{10} / m_{00}$; $\Delta y = m_{01} / m_{00}$

Δ z is not computable by this method. Also, this method cannot give rotational information for objects which have an axis of reflection symmetry (e.g. parallelograms, triangles) as the tensors all go to zero in such cases. So given an RPP whose attitude has to be determined we need to:

a) Check whether any axis of reflection symmetry exists.

b) Check whether it will have any axis of reflection symmetry under any affine transformation.

c) If the face has any axis of reflection symmetry or will have it under affine transformations, subject it to distortion which removes the axes of reflection symmetry.

The procedures for these steps are as follows:

a) Symmetry conditions for polygons: To check a given polygon for axes of reflection symmetry, we use the concept of the *Voronoi diagram.*

4.2.1) Voronoi diagram: Given a set of N points corresponding to the vertices of the polygon, let x^i and x^j be two of the points. Let P(x^i , x^j) be the half plane containing x^i that is defined by the perpendicular bisector of x^i x^j . The intersection of N - 1 such half planes, denoted by V(i) is called the Voronoi polygon associated with p^i . Note that the polygons are unbounded. For N points there are N such polygons which partition the plane into a net called the *Voronoi diagram.* The construction of the Voronoi diagram for a pentagon is shown in Fig. 2.

Let the vertices of a polygon be $x^1, x^2, ..., x^N$ where

$$x^i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

The points v^{ij} such that

$$v^{ij} = \frac{x^i + x^j}{2}$$

will then be the extremities of the Voronoi diagram.

Given a polygon its Voronoi diagram is constructed and the points v^{ij} obtained. For high complexity polygons it is usually computationally more efficient to use this procedure than to directly evaluate v^{ij} from the x^i and x^j. The symmetry conditions are then determined as explained below.

The symmetry conditions for a polygon depend on whether it has an odd or even number of vertices.

(i) If the number of vertices is odd, the axis of symmetry should pass through a vertex and the mid-point of two other vertices.

(ii) If the number of vertices is even, the axis of symmetry passes through two vertices or two mid-points.

For an odd polygon, an axis of symmetry to exist and pass through a point x^k, the required condition is that there exist a set of x^i and x^j ($i \neq k$, $j \neq k$) such that

$$<x^i - x^j, v^{ij} - x^k> = 0$$

such x^i and x^j will form pairs of points symmetric with respect to the axis through x^k.

If there is to be no axis of symmetry through x^k, then

$$<x^i - x^j, v^{ij} - x^k> \neq 0$$

for $i \neq k$, $j \neq k$. This is a necessary and sufficient condition to guarantee the nonexistence of any axis of symmetry through x^k. The same procedure is repeated for all vertices to verify whether the polygon has any axis of reflection symmetry.

For an even vertex polygon, two different kinds of axes of symmetry can exist.

i) Axis passing through two vertices.

ii) Axis passing through two midpoints of vertices.

i) Let x^a and x^b be the vertices to be checked. An axis of symmetry will pass through these vertices if there exists a set of x^i and x^j such that

$$<x^i - x^j, v^{ij} - x^a> = 0$$

and

$$<x^i - x^j, v^{ij} - x^b> = 0$$

for $i \neq a$ or b, $j \neq a$ or b. Such x^i and x^j points will form pairs which are symmetric with respect to the axis joining x^a and x^b. If the line joining x^a and x^b is not to be an axis of symmetry then

$$<x^i - x^j, v^{ij} - x^a> \neq 0$$

or

$$<x^i - x^j, v^{ij} - x^b> \neq 0$$

for $i \neq a$ or b, $j \neq a$ or b. This is a necessary and sufficient condition to guarantee the nonexistence of any axis of symmetry through x^a and x^b . This procedure is repeated for all vertices to verify whether the polygon has any axis of symmetry.

ii) Let v^{pq} and v^{rs} be the vertex midpoints to be checked. An axis of symmetry will pass through these if there exists a set of x^i and x^j such that

$$\langle x^i - x^j, v^{ij} - v^{pq} \rangle = 0$$

and

$$\langle x^i - x^j, v^{ij} - v^{rs} \rangle = 0$$

for i ≠ p, q, r or s and j ≠p, q, r or s. Such x^i and x^j will form pairs of points symmetric with respect to the axis joining v^{pq} and v^{rs}.

If the line joining v^{pq} and v^{rs} is not to be an axis of symmetry then

$$\langle x^i - x^j, v^{ij} - v^{pq} \rangle \neq 0$$

or

$$\langle x^i - x^j, v^{ij} - v^{rs} \rangle \neq 0$$

for i ≠ p, q, r or s and j ≠ p, q, r or s. This is a necessary and sufficient condition to guarantee the nonexistence of any axis of symmetry through v^{pq} and v^{rs} . This procedure is repeated for all combinations of vertices to check that no axis of symmetry exists.

b) Having verified that no axis of symmetry exists for a polygon, we need to further verify whether any axis of symmetry will exist under any affine transformation of the face.

The conditions derived above for no axis of symmetry to exist are of the general form

$$U^T V \neq 0$$

where U and V are two dimensional vectors. Under an affine transformation A, the condition becomes

$$U^T A^T A V \neq 0$$

whether this condition will be satisfied or not depends on the nature of U and V i.e. on the nature of the polygon and also on what kind of affine transformation it is subjected to i.e. on A. Thus a nonsymmetric triangle can be affine transformed to an equilateral or isosceles triangle which has axis of symmetry.

c) If the face has an axis of symmetry as verified in a) or b) then it is subjected to distortion which removes the axes of symmetry. It has been found that while for any particular distortion there always exists an affine transformation that would yield an axis of symmetry, if the polygon is subjected to two separate distortions which are antisymmetric with respect to each other, there exists no affine transformation which yields axes of symmetry for both cases. If these two distortions are referred to as D_1 and D_2 , then the procedure consists of subjecting the polygon to D_1 and checking it according to a) and b) to see whether it has any axis of symmetry. If it does it is subjected to D_2 and by the above argument it will not have any axis of symmetry.

It has been found that polygons with three or four vertices can always be affine transformed to symmetric polygons. So a *minimum of five points* are needed to obtain a nonsymmetric polygon. A technique has been developed whereby it is not necessary for all the five points to physically lie on the polygon. If we have three points on the polygon, the other two points can be obtained as functions of coordinates of these three points. This is shown for a triangle in Fig.3, where P_1 , P_2 and P_3 are points on the triangle (its vertices) and Q_1 and Q_2 are artificially created points. The five points should be positioned such that the polygon formed by them is the distortion D_1 or D_2 referred to above. Two such distortions are shown in Fig.3 b and c.

4.3) Experimental Results: Fig.4 shows certain objects that have been used to test the simulation of the MIAG algorithm. Fig.5 shows the output of the program for recognition and attitude determination of an object under two different orientations. Fig.6, 7 and 8 refer to certain physical objects that have been used to test the MIAG algorithm. These objects are a simple polyhedral structure, a space shuttle model and a space station model. These figures show these objects and their thinned wireframes.

5) CONCLUSION:

The MIAG algorithm has been extended for the attitude determination of general polyhedral objects. The algorithm has been tested under conditions simulating space conditions and the results are presented in this paper. Work is in progress to extend the algorithm to the general case of recognising and localising any general object.

REFERENCES

[1] C.W.Richard and H.Hemani, "Identification of Three-Dimensional Objects using Fourier Descriptors of the Boundary Curve". *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, Mar.1980, pp127-136

[2] S.A.Dudani, K.F.Breeding, and R.B.McGee, "Aircraft Identification by Moment Invariants", *IEEE Trans. on Computers*, vol C-26, no.1, October 1977, pp. 39-45

[3] L.T.Watson and L.G.Shapiro, "Identification of Space Curves from Two-Dimensional Perspective Views", *IEEE Trans. Pattern Anal. and Machine Intell.* vol. PAMI-4, Sep.1982, pp.469-475

[4] D.Marr and T.Poggio, "Cooperative computation of stereo disparity", *Science*, vol.194, pp.283-287, 1977

[5] D.Marr, Vision: *A Computational Investigation into the Human Representation and Processing of Visual Information.* W.H.Freeman and Co., 1982

[6] O.D.Faugeras, M.Hebert, E.Pauchon, J.Pouce, "Object Representation and Positioning from Range Data," *Robotics Research*, First Intrl. Symp., 1984

[7] B.K.P. Horn "Robot Vision", MIT Press, 1986

[8] K.A. Stevens et al., "Understanding Images at MIT: Representative Progress", *Proc. DARPA Image Understanding Workshop*, 1980, pp 15-19

[9] J.R.Kender, "Shape from Texture: A Computational Paradigm", *Proc. DARPA Image Understanding Workshop*, 1979, pp.134-138

[10] A.P.Witkin, "Shape from Contour", Phd dissertation, MIT, Cambridge MA, 1980

[11] H.G.Barrow and J.M.Tenenbaum, "Interpreting line drawings as 3D surfaces" Proc. 1st Annu. Nat. Conf. AI", 1980

[12] Bamieh B. and deFigueiredo R.J.P., "A General Moment Invariants / Attributed-Graph Method for Three-Dimensional Object Recognition from a Single Image", *IEEE Journal of Robotics and Automation*, March 1986, pp 31-41

[13] Pavlidis T., "Algorithms for Graphics and Image Processing", Computer Science Press, 1982.

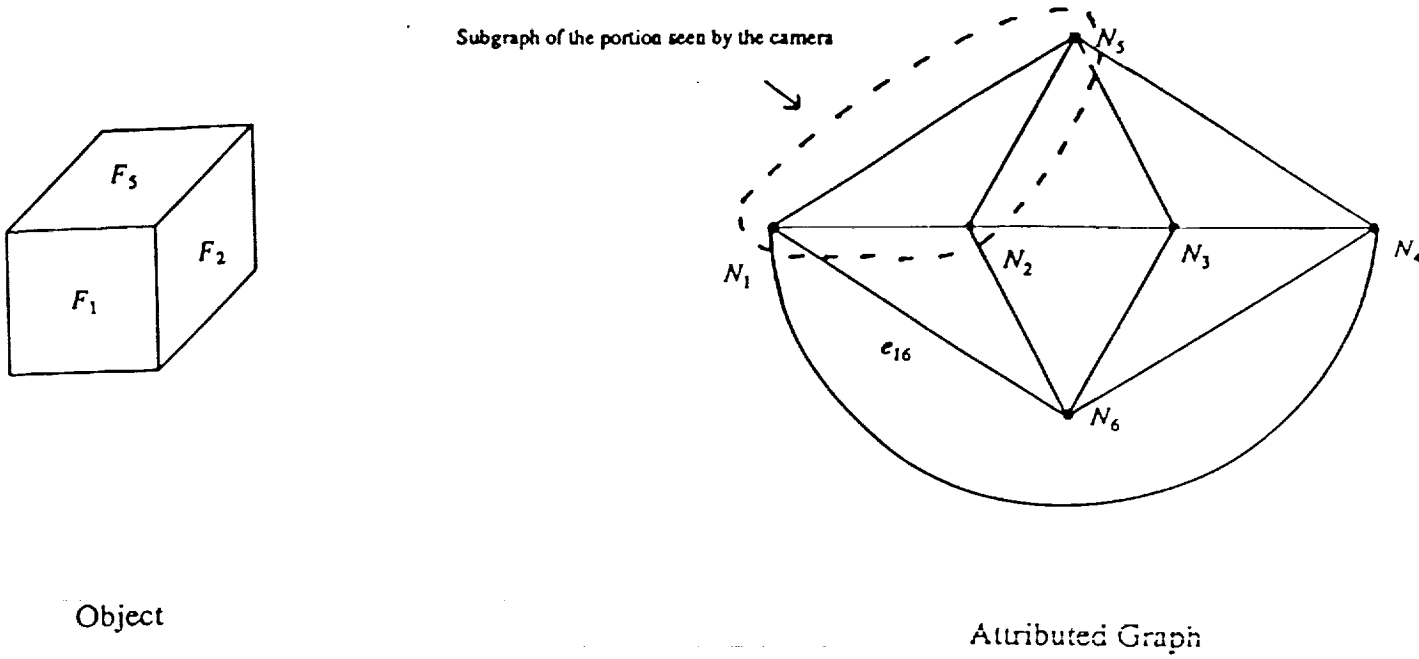


Fig.1 Wireframe of a cube and its attributed graph representation

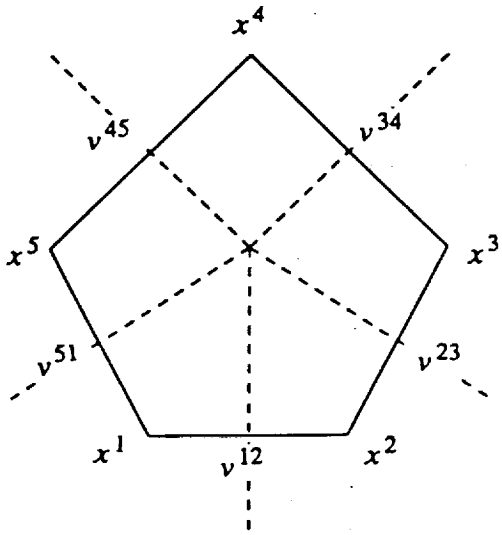
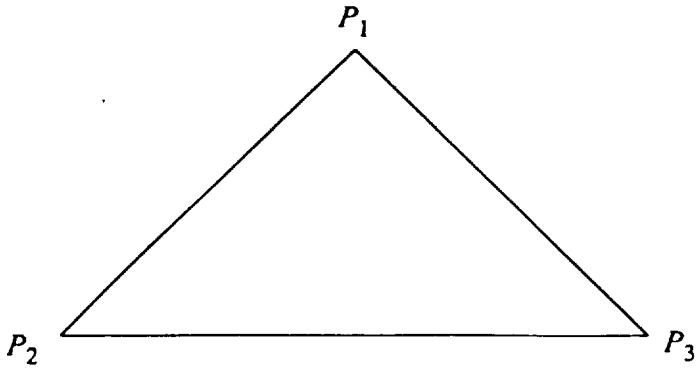
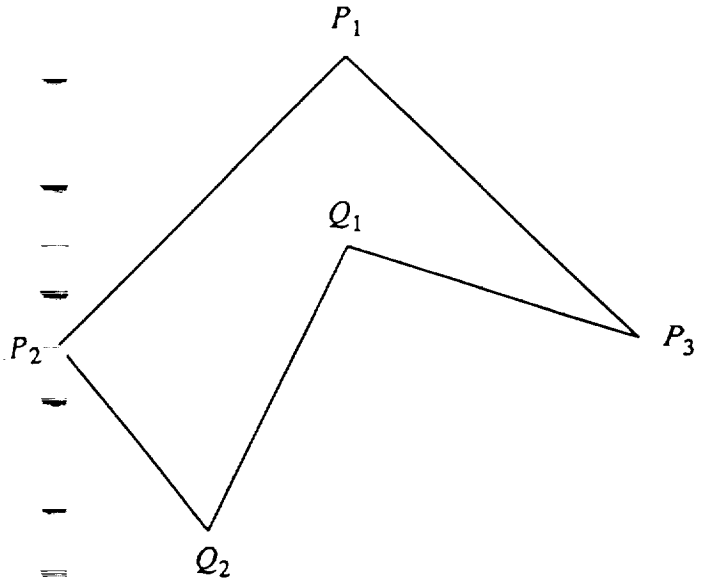


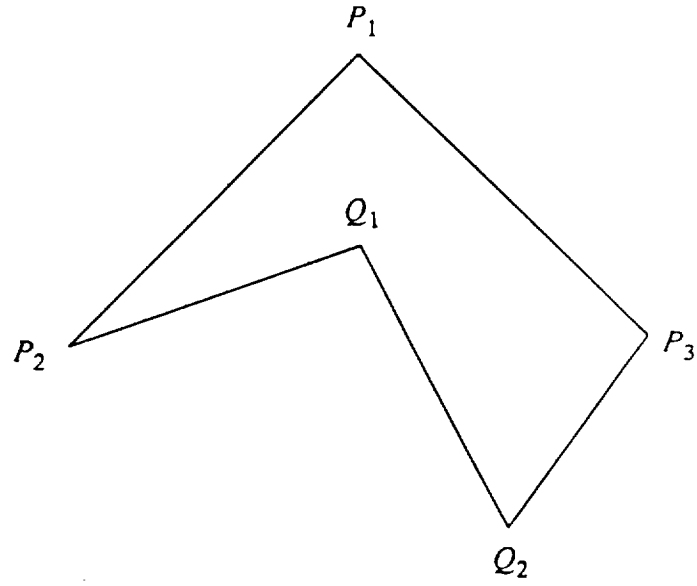
Fig.2 Voronoi diagram of a polygon



a: Triangle

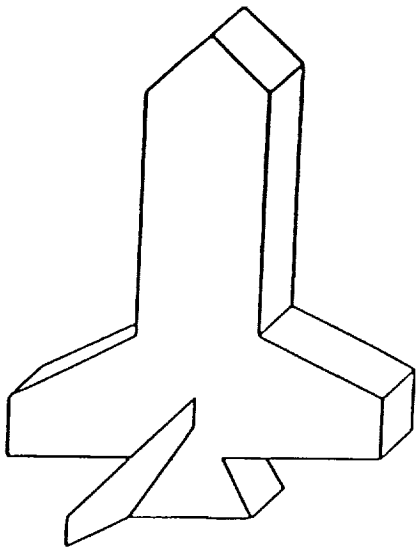


b: D_1

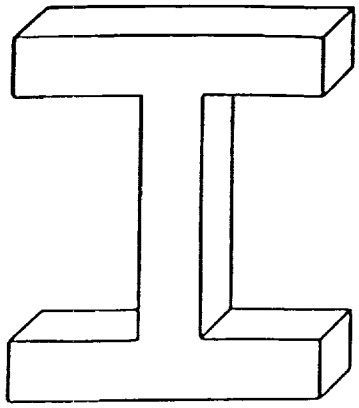


c: D_2

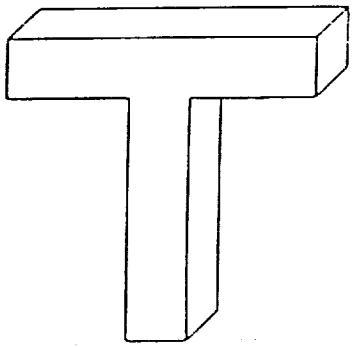
Fig 3: A triangle subjected to D_1 and D_2



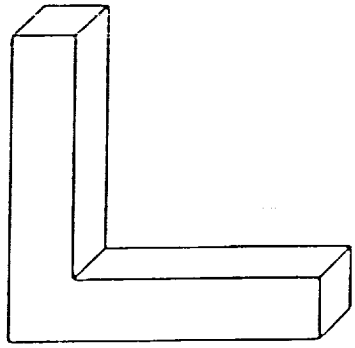
Space Shuttle



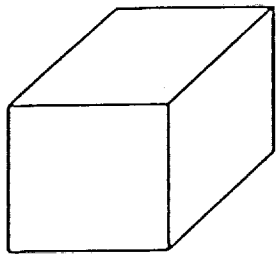
I-Beam



T-Beam



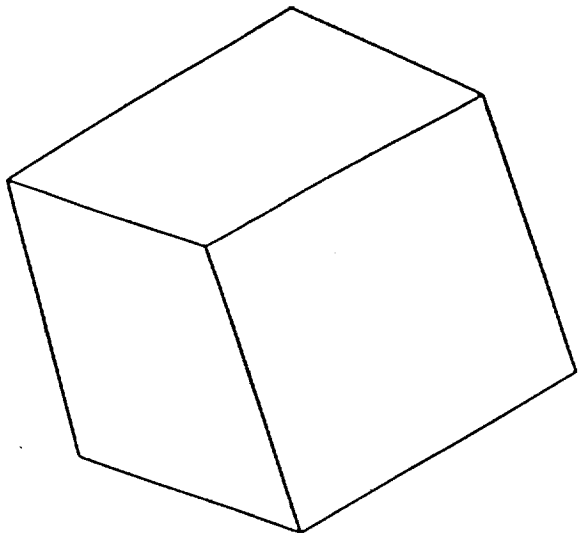
L-Beam



Cube

Fig.4 Examples of objects used to test the MIAG algorithm

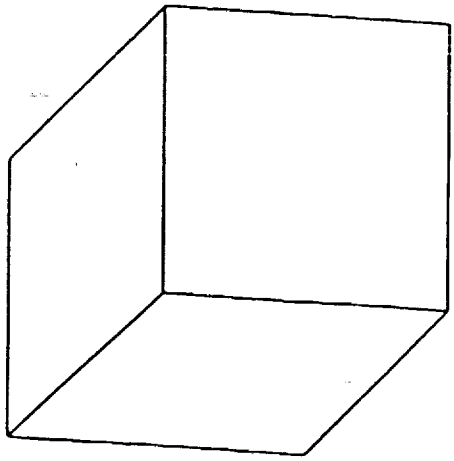
Fig.5 Examples of Recognition and Attitude Determination using MIAG Algorithm



```
roll = 45
pitch = 45
yaw = 45

Matched to cube

Match correspondences are:
Wireframe face # ---> Model face #
0 ---> 0
1 ---> 1
2 ---> 4
```



```
roll = 90
pitch = 30
yaw = 90

Matched to cube

Match correspondences are:
Wireframe face # ---> Model face #
0 ---> 0
1 ---> 1
2 ---> 5
```

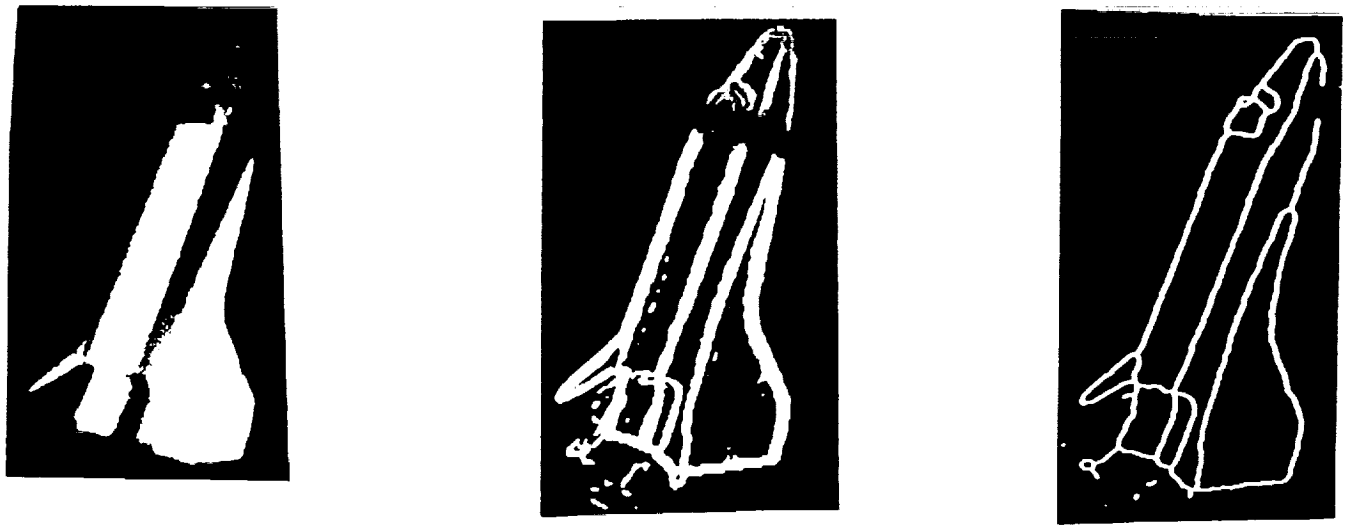


Fig.6 Space Shuttle

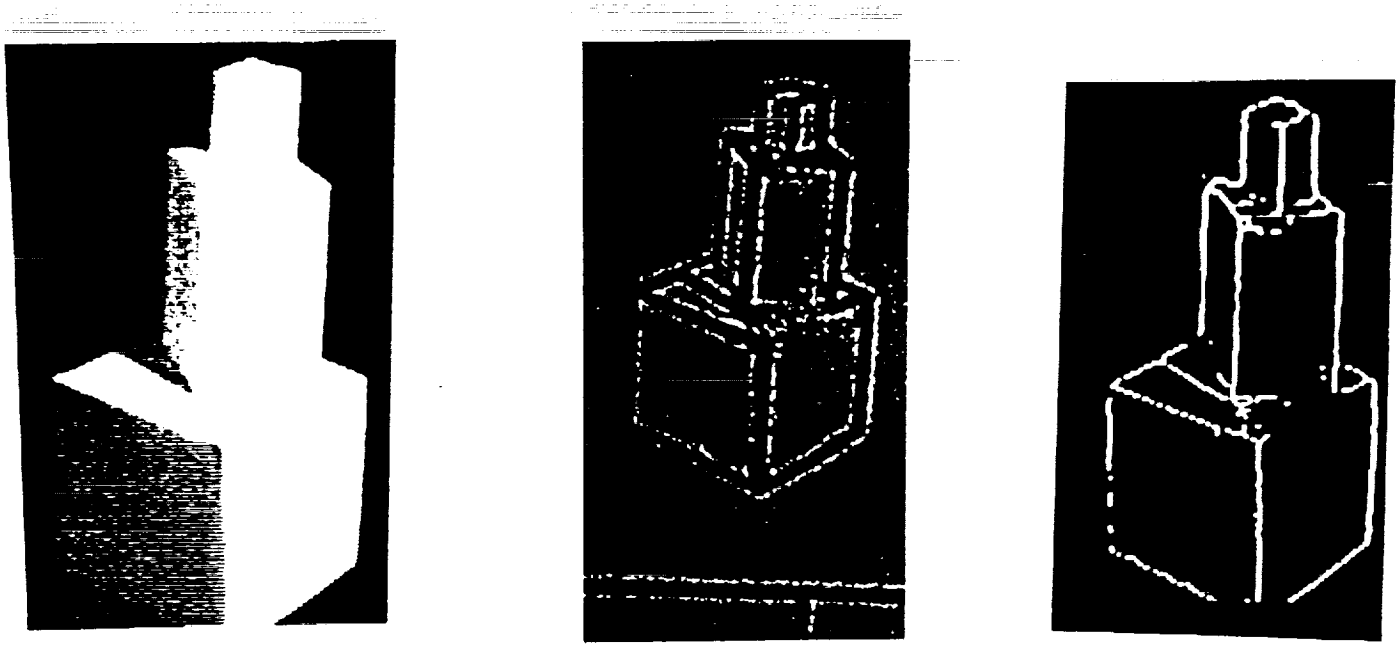
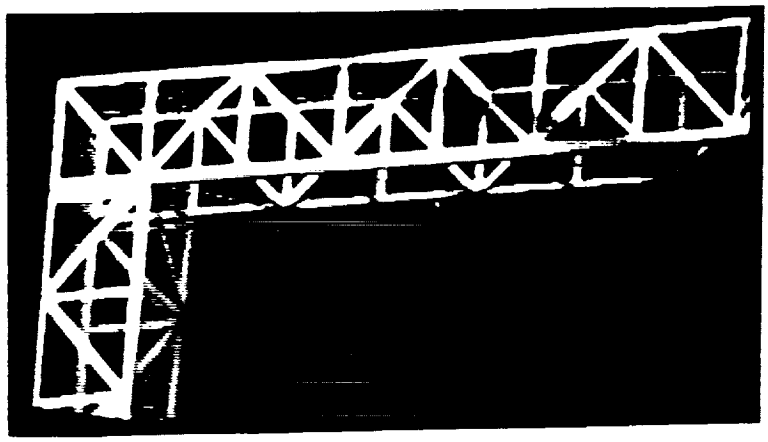


Fig.7 Polyhedral Structure

Object ---> Edge Detector Output ---> Edge Thinning Output



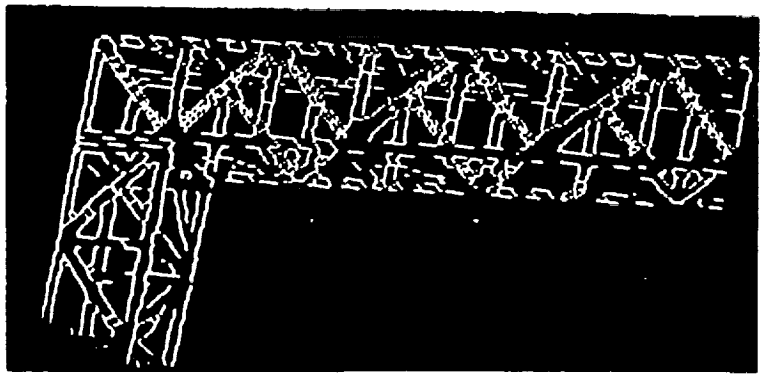


Fig.8 Part of Space Station

Invited Paper from the Proceedings of the 1987 IEEE International Conference on Robotics and Automation, Raleigh, North Carolina, March 30-April 2, 1987.

ROBOTIC VISION/SENSING FOR SPACE APPLICATIONS †

Kumar Krishen * Rui J.P. de Figueiredo ** Olin Graham *

*Tracking and Communications Division, Code EE, NASA/Johnson Space Center, Houston, Texas 77058

**Electrical and Computer Engineering Department, Rice University, Houston, Texas 77251-1892

Abstract

In the past decade, vision/sensing systems techniques have received significant attention for applications in robotics and automation. The thrust of automation and robotics for space applications has been proposed for increased productivity, improved reliability, increased flexibility, higher safety, and for the performance of tasks unsuited to humans. These benefits can be insured through automating time-consuming tasks, increasing productivity/performance of crew-accomplished tasks, and performing tasks beyond the capability of the crew. This paper provides a review of efforts currently in progress at the NASA/Johnson Space Center and at Rice University, the accomplishments to date, and some of the anticipated future developments. Both systems and algorithms are discussed. The evolution of future vision/sensing is projected to include the fusion of multisensors ranging from microwave to optical with multimode capability to include position, attitude, recognition, and motion parameters. The algorithms for information extraction would incorporate aspects of intelligence and knowledge for the interpolation and extrapolation of the needed data. The key features of the overall system design will be small size and weight, fast signal processing, robust algorithms, and accurate parameter determination. These aspects of vision/sensing will also be discussed in this paper.

I. Introduction

The major space program for the next decade for the United States is the NASA permanently manned Space Station. The Space Station will be a multi-purpose facility with missions which will include science and applications, observation, technology development and demonstration, commercial laboratories and production facilities, and operational activities such as servicing/maintenance, repair of satellites, support of unmanned platforms, assembly of large space systems, and as a transportation node for transfer to other orbits or planetary missions. The importance of automation and robotics (A&R) for the establishment and operation of the Space Station has been stressed by the Senate Appropriations Committee (SAC). During 1984, SAC required that NASA establish an Advanced Technology Advisory Committee (ATAC) to identify advanced automation and robotics technologies for the Space Station program. These directives reflected the view of Congress that the Space Station program should not only incorporate advanced A&R, but should also use this opportunity to stimulate national development of these technologies.

† Some of the work reported here was supported by NASA contract NAS 9-17145, the NSF grant DCR-8318514, and the ONR contract N00014-85-K-0152.

The use of A&R for the Space Station program can be viewed in two major areas: *teleoperated/robotic systems* for servicing, maintenance, repairs, and assembly; and *computerized systems* to reduce the manpower requirements of planning, monitoring, diagnosis, and fault recovery of space systems/subsystems. The overall thrust is to increase the productivity through autonomy, increased operational capability, and flexibility. The 1985 ATAC report [1] contains a comprehensive summary of the goals that should be pursued for the initial station (Table 1) and the mature station, projected for year 2010 (Table 2).

The dominant aspects of the technology for A&R are that it be lightweight, small in size, highly reliable, and multifunction. Soft docking with the Space Station can provide an example of the advanced technology needed for robotic control [2]. The station is likely to be flexible and susceptible to oscillations through docking forces. Those experimental or operational systems/processes requiring a high level of stability would be severely impacted. Hence, controlled robotic berthing systems and structures will become essential.

A simplified schematic which shows the functional elements of an automated system is presented in Figure 1. One of the key elements of the system is sensing and perception. Its primary function is to provide information regarding the position of the object in its environment relative to the system's effector. This function involves isolation, description, identification, location, and data transmission. In a broader context a class of object properties which include geometric, mechanical, material, optical, acoustic, electric, magnetic, radioactive, chemical, and weight may be needed. The type, volume, and precision/accuracy of the data needed will depend on the nature of the task to be accomplished.

It should be pointed out that vision is a tool that has broad applications extending beyond robotics/automation. Vision is needed in fixed locations to keep track of location of objects, to pick and place machines/systems, and to monitor performance of nonautomatic systems. Indeed, vision provides intelligence and flexibility in locating automated and nonautomated machines.

The evolution of vision systems seems to have been dictated by the need for quality inspections and productivity improvement. Early vision systems were cameras whose data was fed to a computer. Algorithms in the host processor/computer determined the necessary scene/object parameters which were needed for controllers for various actions. In the 1970's this approach to vision was successful in many industrial applications; however, it was too slow, bulky, and expensive. As the role of robotics began to be widespread, the need for vision systems which are smaller, faster, more flexible, more intelligent, provide accurate images, have high resolution, sense color, and provide location and orientation data became obvious. In the latest versions, intelligence is another feature added to the list.

Table 1. Proposed Goals for Automation and Robotics Applications, Initial Space Station

Automation

- Electrical Power

  Controllers enhanced by expert systems for
  - Load distribution and switching
  - Solar array orientation
  - Trend analysis
  - Fault diagnosis

- Guidance, Navigation, and Control

  Expert system for
  - Station attitude control
  - Experiment pointing
  - Orbital maintenance and reboost
  - Rendezvous navigation
  - Fault diagnosis

- Communication and Tracking

  An executive enhanced by expert systems for
  - Communication scheduling
  - Rendezvous tracking
  - Data rate selection
  - Antenna pointing

- Information and Data Management

  An executive enhanced by expert systems for control of
  - Subsystem statusing
  - Trend analysis
  - Fault diagnosis
  - Redundancy and configuration management
  - Data base management

- Environmental Control and Life Support

  Controllers enhanced by expert systems for
  - Trend analysis
  - Fault diagnosis
  - Crew alarm
  - Station atmosphere monitoring and control
  - Hyperbaric chamber

Robotics

Teleoperation of mobile remote manipulator with collision avoidance

Mobile multiple-arm robot with dextrous manipulators to inspect and exchange orbital replaceable units

Systems designed to be serviced, maintained, and repaired by robots

Table 2. Proposed Goals for Advanced Automation and Robotics Applications, Mature Space Station (2010)

Automation

- Propulsion

  An intelligent controller for
  - Fuel distribution and management
  - Leak detection and evaluation

- Electrical Power

  An autonomous intelligent controller for
  - Power management
  - Fault detection and isolation
  - Maintenance scheduling

- Guidance, Navigation, and Control

  An intelligent controller for
  - Fully automatic rendezvous and docking
  - Space traffic control
  - Remotely piloted vehicles
  - Collision avoidance

- Communication and Tracking

  An intelligent system for
  - Automatic planning
  - Tracking multiple vehicles
  - Scheduling bulk data storage for communications blackouts
  - Detection, identification, and characterization of general targets

- Information and Data Management

  An intelligent system for
  - Fault detection, isolation, and repair
  - Natural language interface with crew

  A data base manager for
  - CAD/CAE bulk data storage facility
  - Retrieval and routing to requestors

- Environmental Control and Life Support

  An intelligent controller for
  - Ensuring fail-safe/fail-operational modes
  - Fault detection and isolation
  - Chemical analysis of air and water
  - Toxic gas analysis

- Habitability

  An intelligent system for
  - Health maintenance
  - Speech interpretation and synthesis
  - Physiological monitoring
  - Automated medical decisions
  - Trend analysis

- Structures and Mechanisms

  Advanced work station
  Intelligent actuators
  Teleoperators

- Orbiting Platforms

  An intelligent system for
  - Process control
  - Maintenance control
  - Planning and trend analysis

Robotics

Inspection, maintenance, refurbishment, and repair
Fuel and materials transfer
Detecting hazardous leaks
Satellite retrieval and servicing

As the new era in robotics dawns, the 3D scene/object description, along with location, orientation, and motion parameters, will be needed. Multisensor capability is deemed essential in order to have the data for the determination of these parameters. Sensor complements may include both microwave and optical with diverse modes. The fusion of these sensors for the mensuration of various parameters in an efficient manner provides by far the greatest challenge for the organization and use of partial or noisy information in a coherent manner. The compression, storage, and transmission of the information associated with multisensor capability require novel algorithms. Finally, both in the single- and multiple-sensor cases, the need for a dramatic increase in processing speed to achieve real-time vision sensing goals will constitute an important consideration in hardware design and development. Current efforts to develop highly parallel multiprocessor arrays, pipelines, and other appropriate network configurations for this purpoose represent a step in this direction.

In this paper, the requirements for data are viewed from the applications standpoint. A review of the systems aspects of vision/sensing is provided. The progress in the area of algorithm development is summarized. Additionally, new concepts in the areas of sensor and algorithm development are elaborated.

II. General Requirements for Space Vision/Sensing Applications

One of the first functions that has been planned for the Space Station is the *assembly* itself. This can involve mating structures, bolting, locking, and forming joints in the structure itself. An option for the assembly of the structure is the use of Shuttle Remote Manipulator System (RMS), controlled either by astronauts during Extra-Vehicular Activity (EVA) or those in the cabin (Figure 2). The initial vision system could be remote TV cameras with the prime function of providing imagery for inspection. As time proceeds, the assembly processes could involve Orbital Maneuvering Vehicle (OMV) with robotic arms and sensors for vision. The parts assembly could involve resistance spot welding, electron beam welding for continuous joint and seal formation, localized light forming, and precision light machining. Such functions will need dramatic improvement in sensing and perception to achieve accurate mechanical control and monitor the performance of the OMV. The versatility of the OMV can be increased by interchanging specialized lightweight end-effectors [2]. Ultimately, in the distant future (year 2020), autonomous robots could be used to assemble detailed structures in space. In these applications noncontact sensing could involve ranging and identification, as well as motion monitoring.

In the areas of *inspection, maintenance,* and *repair,* many tasks must be performed. *Inspection* will include monitoring fatigue, failures, flaws, meteorite damage, structural vibrations, and environmental parameters. The maintenance of a space system can involve the servicing of avionics, thermal system, brakes, landing gear, TV cameras, recorders, printers, actuators, fluid coolant leaks, heaters, thermostats, circuit breakers, switches, power systems, and aging composite materials. Initially, they could be done using Mobile RMS (MRMS) on the Space Station (Figure 2). In the distant future, the inspection/maintenance/repair functions would be highly automated and continuous (Figure 3). This would mean that the robots have to be versatile and adaptive. The vision sensors for these operations should include a full range of sensing which is independent of natural light. Many servicing functions such as refurbishment, resupply, refueling, cleaning, and storing human refuse and debris, and replacement of batteries could be automated with minimum vision, involving only the video sensors. However, most of these tasks require docking/berthing which, when automated, could be made fast and stable. This operation can be made economical by choosing a vision system which is accurate and very fast.

A major thrust for utilization of space environment is *manufacturing*. Many processes have been envisioned for micro-g stable environment, which can yiueld unique products hard or ikmpossible to manufacture on earth. These processes are long duration, need relatively contamination-free environment, can be dangerous to humans, and/or require disturbance-free surroundings. Thus, it is either mandatory or desirable that automated system/robots provide the servicing and maintenance of these facilities. General Electric, in a recent review [3], recommended that the manufacturing facility to able to perform slicing, polishing, cleaning, sawing, separation, ion implant, photo resist, annealing, E-beam direct write, and reactive ion itch. One opportunity is to produce computer chips from hazardous material gallium arsenide for super computer systems. The critical limitation on the vision sensors imposed by the manufacturing task is small size. In most cases the need for vision will arise for human monitoring purposes. Thus, a distributed system spread across the laboratory might be required to cover all areas. For the robotic end-effector, in addition to smallness, a highly accurate and high resolution system will be required.

For the *direct control teleoperator systems*, the primary function of robotic vision is to provide information about the position of the object relative to the system's effector (Figure 4). In a direct controlled teleoperator system, some subfunctions may be allocated to the human. Furthermore, the object identification can be delegated to the human. After the description is given by the human [4], the machine generates a model from the observed video image. This model uses a generic description from the data base. The pattern parameters derived from the video-based model provide the necessary information to generate the end-effector position command (Figure 5).

In a *goal-directed teleoperator system*, reliability and flexibility have to be provided, with which it searches, identifies, and locates parts based on existing data base. The CAD/CAM type information and directives are provided within the teleoperator system. In the absence of the vision system, the required object identification and position data must be entered manually. In a higher level goal-directed teleoperator system, the operator supplies goals. These goals are then broken into tasks by the knowledge based CAD/CAM information regarding the processes involved. Complex tasks are broken into simpler commands which are ezxecuted using known rules and sequences. In such a system, a continuous data flow must be maintained from the vision sensors. The location data should not only identify the object but also the position/motion parameters of the end-effector at each instant of time or successive step in the accomplishment of the task.

In general, for the *autonomous robotic systems*, three levels of information are needed. These levels pertain to the scene/world in which the objects are located, the objects themselves, and specific parts of the objects. In most tasks an envelope of parameters can be preprogrammed into the system. For example, in docking and berthing applications the robotic vision/sensing may be needed within a cone of, say, thirty degrees to a distance of 50 meters. Beyond this spatial zone, a radar system may be used for the tracking/motion monitoring. The levels of information depend on the application involved. As an example in the satellite servicing area, the vision/sensing system may have to provide the necessary information to guide a robot/astrobot to a particular area, say, an antenna feed. The satellite could be rotating and translating simultaneously. Furthermore, the antenna could be gimbaling with a certain motion. In this scenerio the data would not only include a 3D dynamic description of the target/object but also its position and rotational parameters with respect to the satellite. To accomplish this, algorithms are needed for the parameter estimation. The vision/sensing instrumentaiton in this case would not only involve fixed field of view video systems, but laser/millimeter wave radars which could be slaved to the antenna feed motion. Doppler signal

processing at microwave or optical frequencies can be used to sense moving parts within a scene. The implication on the vision/sensing systems is clear; several sensors are needed to complete the basic information needed for an autonomous robot. Clearly, for space applications, the size, speed, and weight parameters are of paramount importance.

Autonomous robot performance depends crucially on the vision capabilities. In certain operations, humans can be surpassed by robots based on the memory and vision. The use of color and polarization in future will become increasingly important, since variations in these parameters are related to certain scene features. General application areas for machine vision are given in Table 3 [5].

Measurement Category	*Sample Application*	*Thrust*
Location	Robotics/Ranging	Where
Identification	Assembly/Inspection	What
Recognition	Sorting/Picking	Which
Inspection	Quality	How Good/Accurate
Gauging	Industrial Control	How Large/Much
Counting	Inventory	How Many
Motion	Ranging	When/Where

Table 3. General Application Categories for Machine Vision

The vision extention to shadowed and occluded regions is important in many applications. The illumination intensity variations along with shadows can, in certain applications, be used to determine object shape, as well as an estimate for the relative motion between the camera and the object. Mathematical models, coupled with real-time imagery, can be used to derive these parameters. Structured multi-spectral lights can be used to derive the 3-D description of the target. Associated with the sensory data is the need for computer architectures [6] which provide high-speed processing, parallel computations/algorithms, associative memories, and intelligence. The transfer and reduction of the sensor data can be facilitated with the implementation of a communications subsystem.

The vacuum of space makes the scattering of light more dependent on the surface geometry and properties. The usual illumination caused by diffraction associated with the reflection of light from objects on earth is negligible. This phenomenon yields pronounced shadows and specular points for space-acquired images. Another detail that should be taken into account is the thermal protection system (e.g., thermal blanket) which obscures direct visual detail of electronic and other systems requiring such protection. In addition to this, occluded areas also exist in the scenes to be surveyed. In view of these environmental constraints, space vision/sensing systems should provide a mode independent of sunlight. To achieve this independence along with penetration, structured optical illumination and microwave sensors might be considered. For this multiple source illumination, the wavelength, intensity, polarization, field of view, and angles of incidence must be chosen for a particular scene parameter (Figure 6). These implementations will ensure appropriate enchancement of texture and dependence on scene color necessary for certain robotic tasks.

The present vision/sensing systems have several serious limitations which result in these being unable to provide data/parameters needed for autonomous robots to have perception/reasoning. These limitations include problems with shadows, occlusions, high contrast, low resolution, 2-D models, rigid bodies, and viewing from a fixed viewpoint. In addition to these, cost-effectiveness, speed, small size, lightweight, high reliability and flexibility, and ease of operation must be considered.

III. Non-Contact Sensors for Data Collection

In robotic vision systems electromagnetic energy emitted or reflected by an object is received, converted/transduced, and finally processed. Transducing usually converts received energy into electrical signals, and processing transforms those signals into parameters needed for perception. This perception then allows the robot to execute a function. Processing can be accomplished in two steps: that of *preprocessing*, which involves compensation/corrections for various hardware characteristics, enhancement, selection, and decoding/decompression; and that of *analysis/interpretation*. The latter step is carried out with specific algorithms to extract the required information.

In the initial vision systems, NASA anticipates the use of stereo televisions for label/feature based object recognition [7]. NASA's television program from Apollo through the Space Shuttle programs has been one of high crew and ground participation and control. The sophistication of work to be performed either by teleoperated or autonomous robots points out several limitations of these systems. These limitations surfaced during the Solar Max Satellite repair, when shadowed surface could not be approached, and the movement and grappling of the Solar Max were severely restricted because of limitations with the manual television camera light level controls. Several modifications are now being developed at the Johnson Space Center (JSC) for future space television systems. These include *predictive auto focusing, programmable predictive scene control with auto zoom, gamma,* and *iris, automated* or *voice control pan, tilt,* and *pointing capability*. Breadboards of these systems are currently under evaluation at JSC [Figures 7, 8]. The development of *illumination systems for enhancing image quality and extracting shape from shading* is also being pursued at JSC [Figure 9]. This development is guided by analytical methods/algorithms discussed in Section IV.

One of the problems in accurate recognition of objects in space is the *Earth background*. Spectrum measurements taken at NASA/JSC [Figure 10] show absorption bands in the near-infrared region. The implementation for the video utilized charge injection devices (CID) to avoid image blooming with optical load compared to charge coupled devices (CCD). The CID sensor (GE TN2505 camera) has adequate responsivity at the 0.94 μm wavelength. This wavelength provides a water absorption notch. Thus, in space this sensor will be earth blind. The calculated attenuation in absorption for a 0.04 μm band is 21.6 dB. The intended use for this sensor in space is pointing and following an object in motion. For this purpose an 8-bit microprocessor is used to control the adaptive system [Figure 11]. This unit has only 2 kilobytes of random access memory (RAM) and 2 kilobytes of read only memory (ROM) which is switchable to 4 kilobytes [8]. The system parameters which can be varied include camera sensitivity, digital filtering, video threshold, and zoom lens setting. The incoming infrared enhanced video is compared to an adjustable threshold, which is varied under the microprocessor control, to generate a serial data stream. An adequate digital filter is used to remove high frequencies from this data stream. The data is then applied to two counters: one counter accumulates the number of pixels above the threshold, which provides a measure of the total area of the object; the other counter is used to count only on the left side of the object in the scene and provides the area on the left side of the scanned area. In a similar way the area is also measured in the top half of the scanned area. These areas are compared to maintain the object in the center of the field of view. This procedure then results in providing commands to the pan and tilt unit. The parameters that affect the dynamic characteristics of the tracking system can be selected for optimum stability and response.

Another advancement in space TV operation incorporated at JSC is the *voice control* [9]. The initial Voice Control System (VCS) has the following performance parameters: (1) isolated and

continuous speech recognition, (2) 200 word vocabulary (at 1.2 sec/word), (3) two person support (3rd optional capability), (4) user-trained (dependent recognition) for high accuracy, and (5) syntaxing capability. This system [Figure 12] will be tested onboard Shuttle and has been made to connect with the Shuttle Audio Distribution System and existing headsets. The system allows complete hands-off control of CCTV functions including: (1) monitor selection, (2) camera selection, and (3) pan, tilt, focus, iris, zoom, and scene track. Future use of voice for TV automation has been projected for the Extra-Vehicular Activity (EVA) astronaut. In this application the astronaut can ask for Heads Up Display (HUD) of vital data from the Shuttle computers. These data can include system parameters, orbital parameters/location, system status, and particular subsystem data [Figure 13]. In the robotics application, these systems can be incorporated using synthetic voice commands.

The need for video system data to be able to interface with digital processors/computers has given impetus to *digital TV technology*. For greater reliability, solid state imagers with sensor sizes 512 × 512 and 1024 × 1024 are being designed. Furthermore, recognition/preprocessing algorithms can be implemented on the same electronics chips making the size of these video imagers small. These technology innovations are also aimed at high resolution systems. As this technology is moving forward rapidly, the need for handling and transmission of high data rates is becoming obvious. For a video system at 5 MHz baseband, an eight bit digitization would generate 80 MBPS data stream. For color TV implementation and multiple systems, this bit rate can multiply significantly. For real time processing of this data, compression techniques have been proposed. These compression algorithms can be hardware implemented using VLSI and VHSIC circuitry. The algorithms must be automatic and transparent to users, not destroy or discard any relevant information, and compress and decompress data at speeds significantly higher than associated device data transfer speeds.

Fourier optical processing offers a method of high speed parallel processing of data needed to support automation and robotics applications [Figure 14]. The inherently parallel nature of optical information processing, coupled with the easy and natural optical Fourier transform and the programmable masks, can obviate numerical processing for many applications. The masks are used to modulate the optical Fourier transform of an input scene [Figure 14]; an optical retransform then allows direct detection of, say, the mathematical correlation between the viewed scene and the reference image, the mask of which is placed at the location of the optical Fourier transform. In this manner any image computation that can be cast in the form of correlation or convolution between object and reference images can be approached with optical processing. Programmable masks are yet in the rudimentary state. Texas Instruments, under joint sponsorship by NASA/JSC and the Army Missile Command, has fabricated a Deformable Mirror Device (DMD) which is under test and further development at this time. The performance degradation caused by optics and diffraction/scattering effects is being studied, and devices/techniques to compensate these effects are being developed. Another area of extention is filtering with a controlled amount of scale and rotation invariance for use in the control/docking applications. Furthermore, Johnson Space Center has designed a high-speed special-purpose geometrical remapping image processor. The device is in manufacture by Texas Instruments; delivery is expected in Fall 1987. Insensitivity to scale and rotation of a viewed object will be the result of one form of mapping, allowing one to use only spatial displacement in tracking the object. The elimination of scale and rotation dependence greatly increases the speed of the vision system for its use in a control network. Experiments with various mappings will result in the design of VLSI cameras whose receptor patterns are best suited to drive a subsequent optical correlator.

For the 3-D vision, laser scanning devices providing range at each point of the target are being explored at NASA/JSC. A matrix of angles/position and ranges within the field of view provides depth/height profile of the object. These measurements can be combined with the video reflectance measurements to yield a 3-D real time visual definition of the target. Two technology implementations of this solid state laser vision device currently available are those using mechanical motion of the mirrored surfaces, and those that involve an inertialess change in the optical properties of a transparent medium. The latter class includes diffraction of light from an acoustically generated periodic structure. Phased-array solid state scanning devices are currently being developed. These devices can provide fast, accurate, and lightweight laser vision. The data from laser vision devices can also be used for automatic zoom/focus control of video systems.

The laser vision measurements are dependent on the intensity of reflected radiation. If coherent radiation is used in order to generate an image of the object information from the amplitudes, as well as the phases of the scattered radiation, a 3-D reconstruction of the object can be made. Such devices are known as holographic devices. The source of coherent radiation can be a solid state laser that can in principle provide a resolution of the order of about 1 μm. Part of the radiated beam is deflected toward the detector [Figure 15], where it interferes with the backscattered light from the object. The hologram can then be generated using known reproduction processes, or a three dimensional description of the object such as a Fast Fourier Transform (FFT). Several applications of holographic scanners have been discussed by Sincerbox [10]. Some of these are directly applicable to space robotics systems.

Microwave systems have been used to detect relative speed of objects and their range in many applications. Their use in space robotics applications is being studied at NASA/JSC. In particular, millimeter wave radars provide attractive performance parameters, in addition to their small size. The possibility of broader beam than laser systems makes these sensors attractive for initial acquisition of moving objects. A radar at 100 GHz has been developed at NASA/JSC [11]. This particular system is for use on a Man Maneuverable Unit (MMU) to provide relative range and velocity to the object [Figure 16]. The radar is designed to operate over the range of speeds from 0.1 to 2.0 fps. This type of radar, operating at several carrier frequencies, can be used to measure backscattering coefficients for various polarization combinations. These coefficients are object structure dependent. There is also possibility of penetration through thermal protection and obscuration caused by nonmetallic objects. This data can be utilized in an iterative manner with that of the video systems to provide scene definition/parameters in certain situations/scenerios.

IV. Algorithms for Information Processing

Robotic vision, as it is generally applied at present, is mainly based on video-acquired data. The acquisition of this data takes place in the form of time sequences of images obtained by one or more vidicon or solid state cameras. In this process, A/D conversion, image arithmetic (e.g., addition/subtraction of two successive images), and formation of pixel space constitute steps usually taken. Both 8-bit gray level and 8x8x8 bit color systems are commonly used. However, the resolution and linearity achieved by the state-of-the-art cameras are somewhat limited. To compensate for this, many cameras are needed in some industrial applications. For example, one sheet metal side rail inspection system uses over three hundred cameras [5]. The corresponding images are patched together into a super pixel space.

Image acquisition is followed by several stages of *information processing*, namely (i) *image preprocessing* (filtering,

restoration, and enhancement), (ii) *feature extraction* (e.g., extraction of centroids, perimeters, and areas of patches in the image, Gaussian curvature from the shading on these patches, min/max distances, edges, corners); (iii) *Object recognition/classification* ; (iv) *estimation of motion parameters* of these objects (position, velocity, attitude, and attitude rate), and (v) *image understanding/scene interpretation*

Active research in vision information processing is being pursued at various centers in this country, and has been widely documented in the literature (see, e.g., [12]-[17], [30] and the references therein). In what follows, we highlight the basic philosophy and thrust of some of the algorithmic work currently in progress at Rice University. This work is being carried out at the Computer Vision Research Laboratory. The equipment used in this research includes a system of solid state cameras such as the Hitachi KP232 with accessories, connected to Series 100 Imaging Technology circuit boards, and from these to a SUN Microsystems 3/160 color workstation and a VAX 11/750 minicomputer with several peripherals. Cardboard and plastic models of space objects, such as satellites, the shuttle, and parts of the space station (see Fig.17) are being used in a studio environment, under controlled illumination conditions, for generating the laboratory data needed to provide insights for algorithmic development. The final testing of the developed algorithms will be carried out on the computer vision testbed (Fig. 9) at the NASA/Johnson Space Center mentioned earlier.

The main points of view on which the algorithmic work at Rice is based as well as some of the developing results are now listed below:

(i) *Model-Based Approach:* In most space scenarios as well as in industrial environments, the objects being viewed are unlabeled objects from a library of known and precisely defined objects. So in our algorithmic work, we assume this to be the case, and attempt to capitalize on it to reduce the hardware requirements and the computational effort needed to execute many of the 3D vision functions executed by a robot, as indicated in (ii),(v) and (vi) below.

(ii) *The FIAG and MIAG Representations of 3D Objects*: We define a rigid 3D object η by the surface $S(\eta)$ bounding it. We assume $S(\eta)$ to be piecewise smooth and satisfying conditions which permit it to be represented as a union of smooth surface patches or "faces" $F_1, F_2, ..., F_n$. (See Fig. 18 .) The boundary of each face F_i consists of a closed contour C_i, the union $\bigcup_{i=1}^{n} C_i$ of all such contours constituting what we call the *"3D wireframe"* of the object η. This allows the symbolic representation of the object η by an attributed graph $G(\eta)$ (which wraps around $S(\eta)$), the nodes N_i, $i = 1, ..., n$, and the links e_{ij}, $i = i$, $j = 1, ..., n$, of $G(\eta)$ the connectivity between adjoining faces F_i and F_j. Let an m-vector $I^i = col\,(I^i_1, ..., I^i_m)$ represent a set of attributes (features) associated with the face F_i which are *invariant* under 3D translation, scaling, and rotation. Examples of such I^i are the sets of numbers expressing the Gaussian curvature or mean curvature of F_i. We call the attributed graph $\vec{G}(\eta)$ obtained from $G(\eta)$ by assigning to the nodes N_i, $i = 1, ..., n$, respectively, the attribute vectors I^i, the FIAG (Feature-Invariants/Attributed-Graph) representation of the object η. A special case of this representation is the MIAG (Moment-Invariants/Attributed-Graph) representation of polyhedral objects proposed in [18], in which I^i constitutes a set of 2D moment invariants of F_i (these being invariant with regard to 3D translation, scaling, and translation). (See Figure 19).

(iii) *Wireframe Extraction*: For the utilization of the FIAG and MIAG representations (see (v) and (vi)), it is necessary to extract the *2D-wireframe* $W(f(\eta)) = W(f)$ of the image f of η. $W(\eta)$ is made up of the set of edges in f. In order to extract $W(\eta)$, a number of operations have to be performed on the raw image data. These are Gaussian 2D filtering (to remove noise), edge detection by application of a (3 × 3) Sobel operator followed by thresholding (horizontal and vertical), median filtering (to delete spurious edges), edge thinning, and model-based reconstruction to recover missing segments in $W(\eta)$.

(iv) *Wireframe Labeling and Its Attributed Graph Representation:* $W(\eta)$ is labeled by an algorithm [19] in which each mesh of M_i of $W(\eta)$ is traversed clockwise and the outer loop counterclockwise. This algorithm also determines the adjacency of the regions R_i enclosed by the meshes M_i. The regions R_i defined by $W(\eta)$ correspond respectively to the visible faces F_i on $S(\eta)$. Also, in the polyhedral case, the moment invariants of R_i are the same as those of F_i. Thus, in this case, a MIAG representation $\vec{G}(f)$ can be obtained for f, based on $W(f)$, and clearly $\vec{G}(f)$ is a subgraph of $\vec{G}(\eta)$.

(v) *3D Object Identification from a Single Image*: It follows from the preceding that the identification of an unlabeled object η from its image f is achieved by obtaining $\vec{G}(f)$ and matching $\vec{G}(f)$ to a subgraph of one of the graphs $\vec{G}(\eta^i)$, $i = 1, ..., N$, in a library of models $\{\eta^i\}$ by a subgraph isomorphism algorithm [18].

(vi) *Motion Parameter Estimation with a Single Camera*: Using appropriate camera calibration, all the motion parameters, except for a scaling factor, can be measured by means of a single high precision camera. For this purpose, there are basically two model-based methods available: One, based on the contraction of the moment tensors of a surface patch of the model and its image, determines the attitudes vector $\underline{\theta}$ (raw, pitch, and yaw) and attitude velocity $\dot{\underline{\theta}}$. (See [18] and [20] for details.) The other, based on the correspondence (assumed known) of eight points on the image f and eight points on the model η^i (assumed located and oriented in a standard position), yields all the motion parameters (position, velocity, attitude, and attitude vector) except for a scaling parameter. This second method has been extensively discussed by Longuet-Higgins[21], Huang[22], Haralick[23] and others. Our contribution to the second approach is the ability to establish the point correspondence needed for it by obtaining $\vec{G}(f)$ from f and matching $\vec{G}(f)$ to a subgraph of $\vec{G}(\eta)$ along the lines indicated earlier.

(vi) *Application to Stereovision*: The approach just mentioned can also be used to establish the point correspondence needed for stereovision.

The model-based approach proposed in the above paragraphs, even though not fully investigated, appears to offer two advantages: First, it reduces the hardware by using only a single camera for identification and for most of the motion parameter estimation of a moving object. Secondly, the attributed graph representation and subgraph matching proposed here avoids the computational effort and problems inherent in the probablistic search associated with the matching at the feature level [30].

(vii) Illumination: Shape from shading algorithms are expected to play a significant role in the analysis of space scenes because light scattering from space objects strongly depends on their surface due to the prevailing vacuum. The characteristic strip expansion methods [16,24] have several shortcomings, including sensitivity to measurement noise and a tendency of adjacent characteristic strips to cross over each other, due to accumulation of small numerical errors. Finally, the procedure is not amenable to implementation in parallel form. The variational method [24] that uses an object's occluding boundaries as cues to the recovery of its shape from shading alleviates these limitations. The blending of concepts from variational calculus with those from the best approximation theory can lead to spline-based solutions for the gradient functions determining the local surface shape orientation, as obtained in [25].

(viii) Surface Reconstruction: Algorithms for surface reconstruction from sparse range data can be of value in providing vision to teleoperators. The reconstruction of a surface can also be used as an intermediate step in extracting surface features for machine recognition. Mathematical analysis [26], differential geometry in particular, as well as formal language theory (in syntactic/semantic shape description) can play a significant role in surface reconstruction algorithms as indicated in [27].

(ix) Sensor Fusion: The fusion of data from more than one type of sensors can be of value when any one sensor cannot capture the full information needed. Algorithms for fusion of microwave and video data introduce significantly the wave equation scattering theory into robot vision [28].

(x) High level processing: Apart from techniques based on geometric models, which have been discussed in the previous section, current efforts on machine vision explore AI techniques. These are based on symbolic scene description. Such a description enables one to transform the image into knowledge which can be utilized in the conditioning of the action and the behavior of a robot. Thus scene understanding can be viewed as a reasoning process based on visual data. This process can consist of *common sense reasoning*, for which the knowledge is wide and shallow, and *expert reasoning*, based on the narrow and deep knowledge of experts. A recent effort has been directed toward a frame-based approach to model common sense reasoning [29]. Object-oriented programming is particularly suited for implementing such an approach. In addition, parallelizing the frame manipulation in image understanding appears to be promising for a real-time solution and it is under investigation.

V. *Proposed Future Developments*

As was mentioned earlier, the interaction of natural light with the objects in space has to be accounted for in the algorithms. Furthermore, methods for artificial lighting have to be developed which can provide structured (known distribution) light across the object. The pronounced shadows and specular points due to the vacuum and smooth parts of the object, provide a large dynamic range of the reflected/scattered signal. The intensity changes can be in the 10^8 range. The addition of artificial illumination provides the opportunity to control intensity, wavelength, polarization, and orientation with attendant advantages of increased recognition and shape estimation capabilities. Additionally, color will be another discriminant involved in the algorithms. Analytical studies in these areas should lead to the design of illumination systems for space applications.

The use of laser vision and microwave scattering instruments creates another area of future development. Fast scanning/holographic lasers provide a depth perception of objects, which is quite complex. This depth data can be utilized to iteratively provide a 3-D image of the object by weighting video-acquired image data appropriately. These weights will depend on the surface curvatures as they project in the incidence direction of the laser beam. Both empirical and analytical studies are needed. The microwave backscattering can provide another independent set of data. The shape of certain objects can be directly deduced from this data. In many inspection tasks in which a nonmetallic shielding has obscured the view, such sensing will be mandatory. In other situations the microwave data can be iteratively used with that of TV to arrive at a more definitive description of the object. At the expense of complexity, doppler processing of microwave data can yield the shape of a distant object. The advantages of such a vision are that it is independent of sunlight and it provides a direct measure of range and relative velocity of the object.

Another area of endeavor should be near-field sensing. In this mode a microwave sensor can be very close to the object. A sharp pulse transmitted yields a unique description of the object. This time domain reflectometery is evolving rapidly. Another mode of the system can utilize reflectivity data in the near-field. These techniques have not been explored for the robotic vision applications.

Finally, further research and development is needed in the area of multisensor coordination and fusion. The recognition algorithms are to be extended to include interrelating data from several cameras, laser scanners/holographic systems, and microwave sensors. These algorithms should include motion, rotation, and object changes as functions of space and time. In addition to this, "environmental" data, pertaining to the events/objects and their status, has to be included. These aspects, along with rational models, incorporate artificial intelligence techniques in the scene analysis, driven by common-sense systems and expert systems.

VI. *Conclusions*

This paper is aimed at providing a review of some of the efforts in progress at NASA/JSC and Rice University. The design and development of a vision system for space applications needs several considerations which make them different compared to those used in ground applications. The concerns for space-unique vision systems have been elaborated. Several efforts which need to be undertaken have been discussed. Considerable work has to be accomplished in order to provide robust, lightweight, real-time (fast), and small size vision systems for specific space applications.

VII. *Acknowledgment*

The authors wish to thank Mr. R.S. Sawyer of the NASA/Johnson Space Center for his encouragement and continued support of the research reported in this paper.

VIII. *References*

1. "Advancing Automation and Robotics Technology for the Space Station and for the U.S. Economy," NASA Technical Memorandum 87566, Vol. I, Executive Overview, Vol. II, A Technical Report, March 1985.
2. Tesar, Delbert, "An Assessment of the Development and Application Potential for Robots to Support Space Station Operations," College of Engineering, The University of Texas, September 1985.
3. "Space Station Automation Study, Automation Requirements ...." General Electric Company Report for NASA Contract No. NAS 5-25182, July 1984.
4. Meintel, A.J., and Ralph W. Will, "Space Teleoperation Research," NASA-Langley Research Center Report presented at the American Nuclear Society Executive Conference, Pine Mountain, Georgia, April 21-24, 1985.
5. Murray, Lawrence A., "The Present and Future State of Machine Vision Industry," NASA Automation and Robotics Workshop, Johnson Space Center, January 21, 1986.
6. Cliffone, Donald L., "Sensing for Automation and Robotics," A White Paper for the NASA/OAST Sensor Working Group, August 1985.
7. Holcomb, L.B., R. Larsen, and M. Montemerlo, "The NASA Automation and Robotics Research Program," AIAA/NASA Symposium on Automation, Robotics, and Advanced Computing for the National Space Program, Washington, D.C., September 4-5, 1985.
8. Lipoma, P.C. and D.P. Walton, "A Two-Dimensional Near-Infrared Tracking System," SPIE's Cambridge Symposium on Optical and Electro-Optical Engineering, Cambridge, Mass., September 15-20, 1985.

9. Jordan, William T., "Voice Controlled Closed Circuit Television for the Space Shuttle Orbiter," The Official Proceedings of Speech Tech. '85, New York, April 22-24, 1985.
10. Marshall, Gerald F. (Ed.), Laser Beam Scanning, Marcel Dekker, Inc., New York, 1985.
11. Lichtenberg, C.L., Man Maneuverable Unit Millimeter-Wave (100 GHz) Developmental Radar – Design and Test Results Report, NASA/Johnson Space Center, Houston, Texas, 1986.
12. Rosenfeld, A., and A. Kak, Digital Picture Processing, Vols. I & II, Academic Press, NY, 1982.
13. Marr, D., Vision, W.H. Freeman Press, San Francisco, CA, 1982.
14. Cohen, P.R., and A. Feigenbaum, The Handbook of Artificial Intelligence, Vol. 3, Herristech Press, Stanford, CA, 1982.
15. Ballard, D.H., and C.M. Brown, Computer Vision, Prentice-Hall, Englewood Cliffs, NJ, 1982.
16. Horn, B.K.P., Robot Vision, McGraw-Hill, NY, 1986.
17. Winston, P., Artificial Intelligence, Addison-Wesley, Reading, Mass., 1984.
18. Bamieh, B., and R.J.P. de Figueiredo, "A General Moment-Invariants/Attributed-Graph Method for 3D Object Recognition from a Single Image," IEEE Journal of Robotics and Automation, vol. RA-2, no. 1, pp. 31-41, March 1986.
19. de Figueiredo, R.J.P., B.A. Bamieh, S. Fotedar, E. Hack, K. Trahan, and C. Wu, "Demonstration of a Methodology for Machine Recognition and Attitude Determination of a 3D Object from a Single TV Picture Frame", Technical Report EE8520, Dept. of Electrical and Computer Engineering, Rice University, December 1985.
20. Markandey,V., H. Tagare, and R.J.P. de Figueiredo, "A Technique for 3D Robot Vision for Space Applications", Proc. of Space Telerobotics Workshop, JPL, Pasadena, CA, Jan 20-22, 1987.
21. Longuet-Higgins,H.C., " A Computer Algorithm for Reconstructing a Scene from Two Projections," Nature, vol. 293, pp. 133-135, Sept'81.
22. Tsai,R.Y., and T. S. Huang," Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces," IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-6, No. 1, January 1984, pp. 13-27.
23. Zhuang,X., T. S. Huang and R. M. Haralick," From Two-View Motion Equations to Three-Dimensional Motion Parameters and Surface Structure : A Direct and Stable Algorithm," Proc. IEEE Int. Conf. on Robotics and Automation, San Francisco, California, April 7-10, 1986, pp. 621-626.
24. Horn, B.K.P. and M.J.Brooks, "The Variational Approach to Shape from Shading," MIT AI Laboratory Memo 813, March 1985.
25. de Figueiredo, R.J.P. and V. Markandey, "Recovering Shape of Space Objects from the Shadowing Caused by Illumination," Tech. Report EE-8608, ECE, Rice University, September 1986.
26. Kehtarnavaz, N. and R.J.P. de Figueiredo, "A Novel Surface Reconstruction Framework from 3D Contours," Proc. of SPIE's Cambridge Symposium on Advances in Intelligent Robotic Systems, Cambridge, MA, October 1986.
27. Gregory, J.A., ed. "The Mathematics of Surfaces", The Institute of Mathematics and its Applications Conference Series, September 1984.
28. Shaw, S. W., K. Krishen, and R. J. P. de Figueiredo, "Microwave and Video Sensor Fusion for Shape Extraction of Space Objects" Rice University Report EE-8706.
29. de Figueiredo, R.J.P. and Kuanshen H. Wang, "An 'Evolving Frame' Approach to Learning with Application to Adaptive Navigation", 1986 International Conference on Systems, Man, and Cybernetics, Atlanta, GA, October 1986.
30. Wilcox, B., D.B.Gennery, B.Bon, and T.Litwin "Real-Time Model-Based Vision System for Object Acquisition and Tracking", SPIE International Conference, Los Angeles, 1987.

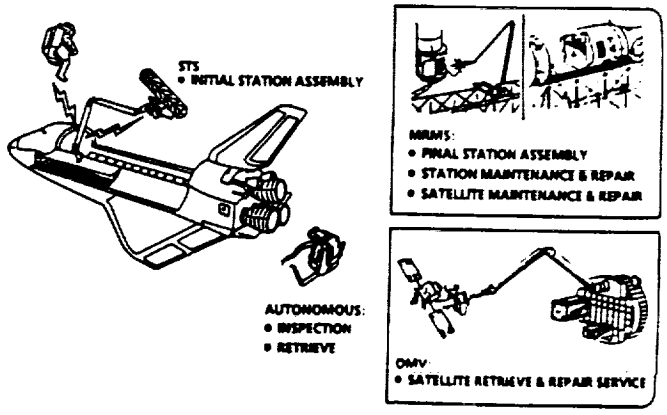


Fig.2 Conceptual Space Station Automation/Robotics

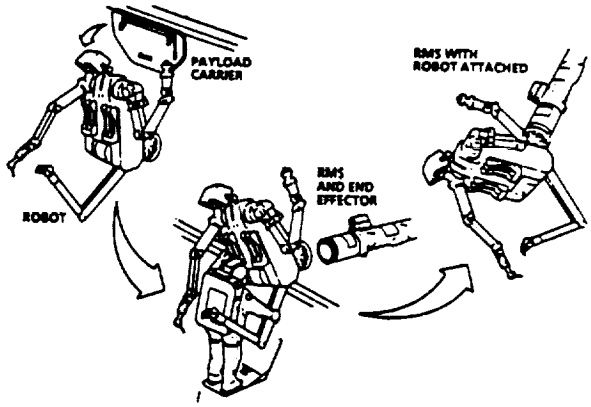


Fig.3 Advanced Space Automation/Robotics Concepts

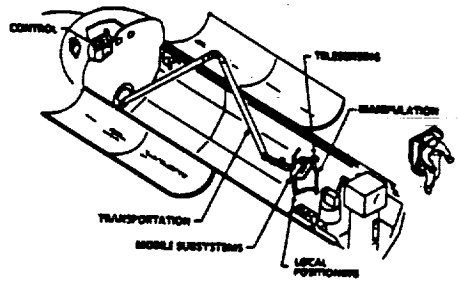


Fig.4 Teleoperator Concepts for Shuttle Operations

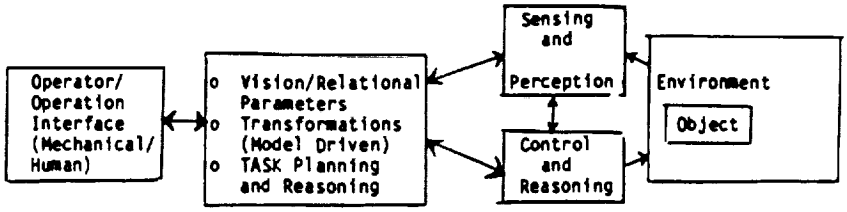


Fig.1 Functional Elements of Teleoperator/Autonomous System

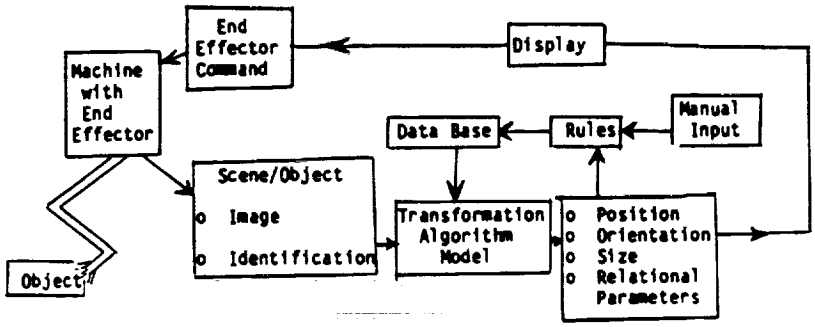


Fig.5 Direct Controlled Teleoperator Vision

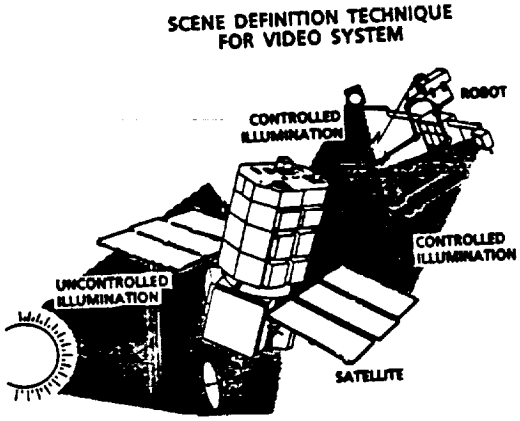


Fig.6 Structured Illumination for Robotic Vision/Sensing

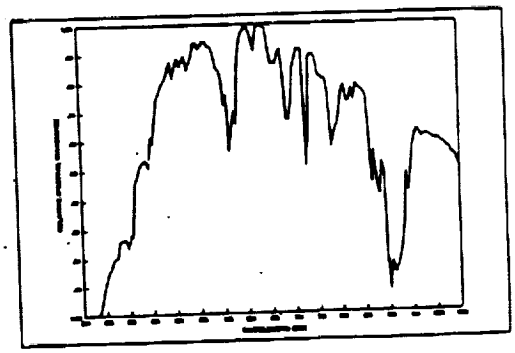


Fig.10 Spectrum analysis of energy reaching the Earth

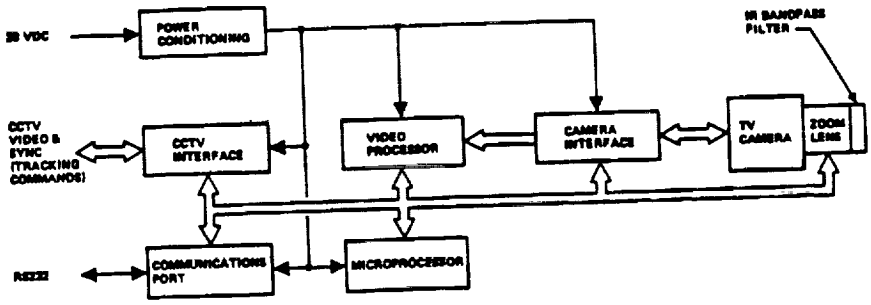


Fig.11 Video Pointing System

Fig.7 Photograph of Smart Television Testbed

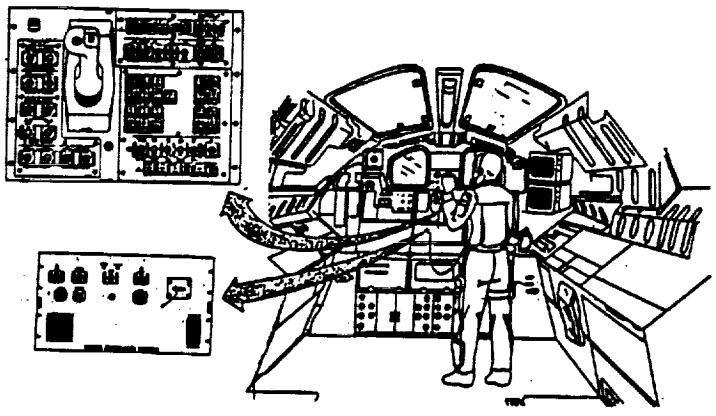


Fig.12 Voice Control System for Shuttle TV

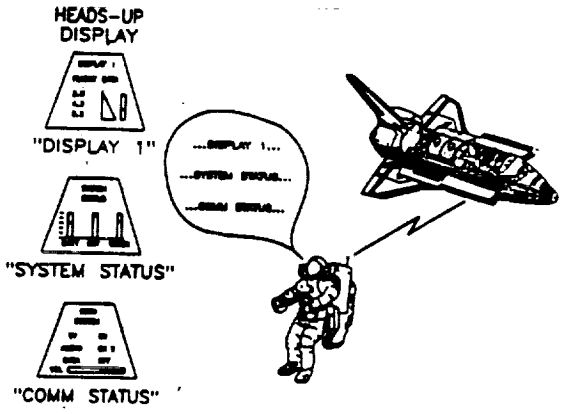


Fig.13 Conceptual EVA TV Voice Command System

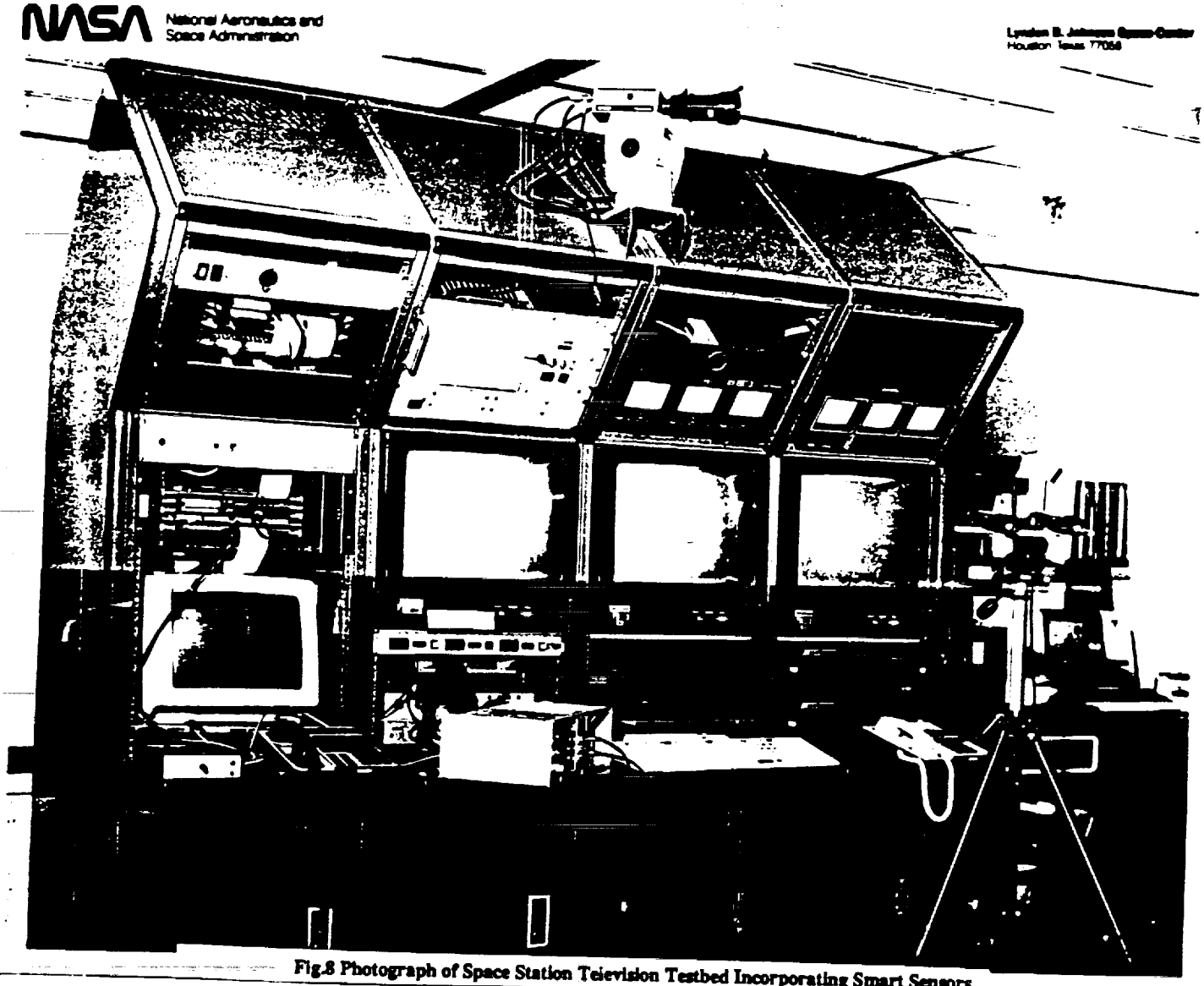


Fig.8 Photograph of Space Station Television Testbed Incorporating Smart Sensors

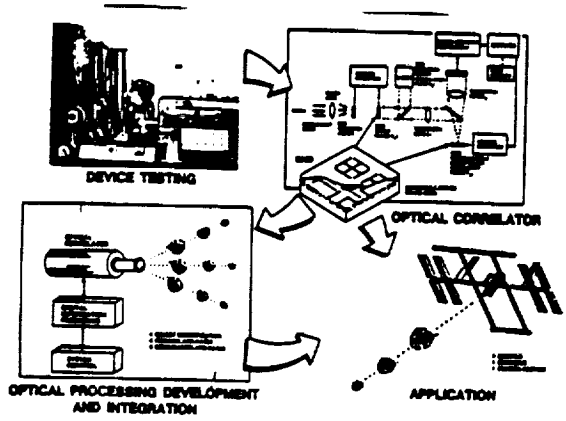


Fig.14 Programmable Mask Technology

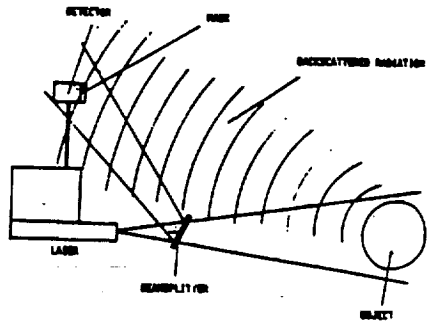


Fig.15 Possible scheme for 3-D image processing based on holography

Fig.9 Testbed for Shape from Shading Experimentation

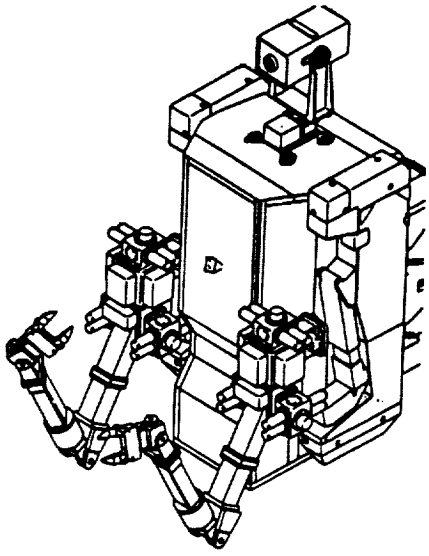


Fig.16 Automated Manned Maneuvering Unit (MMU) with MMU Radar

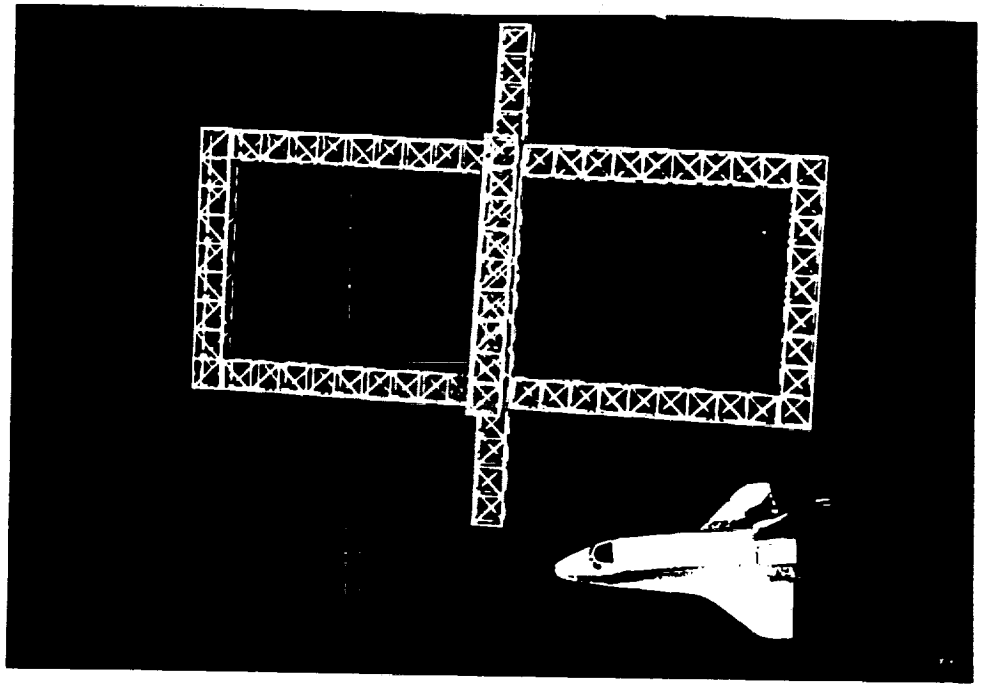


Fig.17 Laboratory models of the Space Shuttle and part of the Space Station

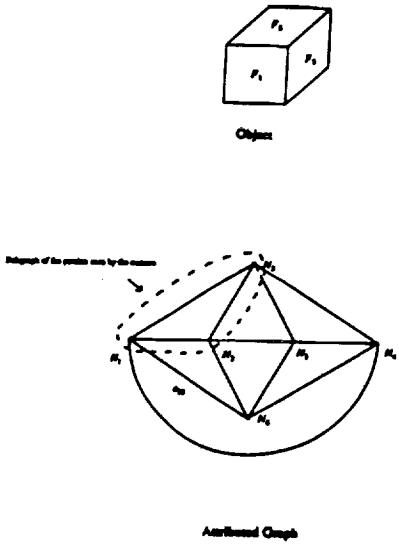


Fig.18 Wireframe of a cube and its attributed graph representation

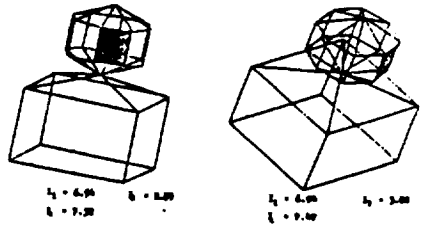


Fig.19 Moment Invariants for one face of the simulated object